

# Practical Exam: Hotel Operations

LuxurStay Hotels is a major, international chain of hotels. They offer hotels for both business and leisure travellers in major cities across the world. The chain prides themselves on the level of customer service that they offer.

However, the management has been receiving complaints about slow room service in some hotel branches. As these complaints are impacting the customer satisfaction rates, it has become a serious issue. Recent data shows that customer satisfaction has dropped from the 4.5 rating that they expect.

You are working with the Head of Operations to identify possible causes and hotel branches with the worst problems.

## Data

The following schema diagram shows the tables available. You have only been provided with data where customers provided a feedback rating.



Before you can start any analysis, you need to confirm that the data is accurate and reflects what you expect to see.

It is known that there are some issues with the `branch` table, and the data team have provided the following data description.

Write a query to return data matching this description, including identifying and cleaning all invalid values. You must match all column names and description criteria. Your output should be a DataFrame named 'clean\_branch\_data'.

Column Name	Criteria
id	Nominal. The unique identifier of the hotel. Missing values are not possible due to the database structure.
location	Nominal. The location of the particular hotel. One of four possible values, 'EMEA', 'NA', 'LATAM' and 'APAC'. Missing values should be replaced with "Unknown".
total_rooms	Discrete. The total number of rooms in the hotel. Must be a positive integer between 1 and 400. Missing values should be replaced with the default number of rooms, 100.
staff_count	Discrete. The number of staff employed in the hotel service department. Missing values should be replaced with the total_rooms multiplied by 1.5.
opening_date	Discrete. The year in which the hotel opened. This can be any value between 2000 and 2023. Missing values should be replaced with 2023.
target_guests	Nominal. The primary type of guest that is expected to use the hotel. Can be one of 'Leisure' or 'Business'. Missing values should be replaced with 'Leisure'.